


Colorization

Deep Learning Project

20182786 강민수
20182836 황태균
20192792 이예진

Contents

- 
- 1. Colorization 이란?
 - 2. Colorization의 한계
 - 3. 주제 선택
 - 4. SOTA 모델 (ChromaGAN, DeOldify, InstColor)
 - 5. 실험 방식
 - 6. Dataset
 - 7. 평가 Metric

Colorization 이란?

흑백 사진을 컬러 사진으로 Deep Learning을 통해 바꾸는 방법론

단색 사진이나 비디오에 컬러 정보를 추가하는 과정

고유한 솔루션이 없는 3차원 색의 값, 밝기를 이미지와 매핑해야 함

예시



Colorization의 한계

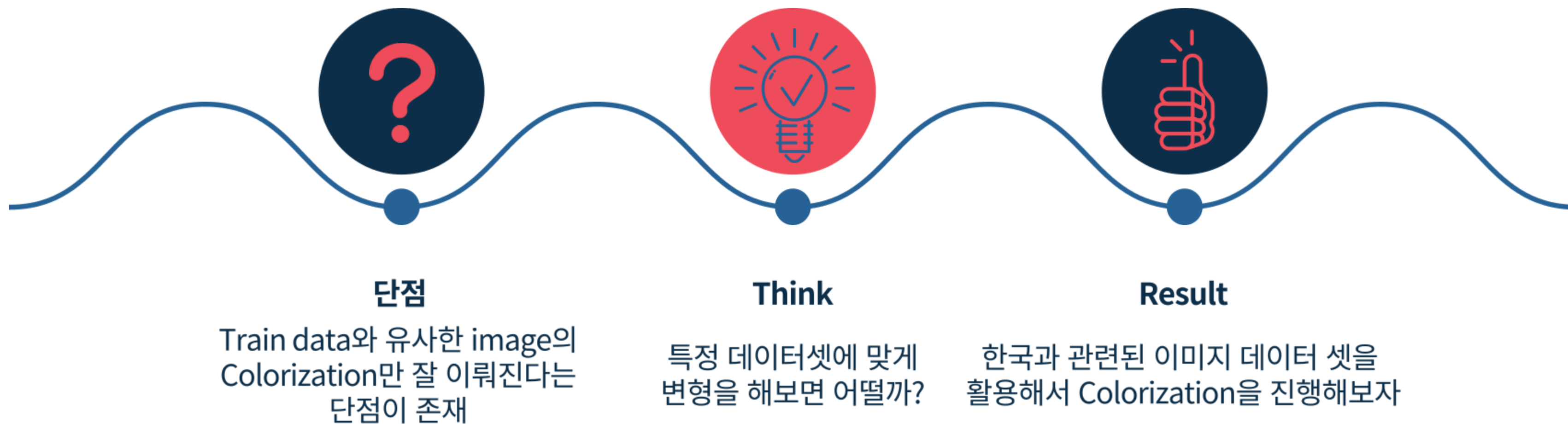
Data-driven task이기 때문에 train data와 유사한 이미지에서만 Colorization이 잘 이루어짐

같은 사물이어도 색상이 어떤 것인지는 모호



1962년의 서울 버스정류장 사진의 색상을 복원한 결과. 순서대로 원본, 색상 복원, 화질 복원+색상 복원

주제 선택



한국의 이미지 데이터셋에 Colorization 적용

SOTA(State-of-the-Art)

ChromaGAN

Paper : ChromaGAN: Adversarial Picture Colorization with Semantic Class Distribution

의미있는 클래스 분포 학습을 위해 자연 색상 이미지의 확률 분포를 학습하고 색상 속성을 생성하기 위해 Generative adversarial networks(GANs)의 강점을 결합하여 학습하는 방식을 사용

이미지를 의미있게 이해하여 흑백사진에 지각적으로 색채를 입힐 수 있음

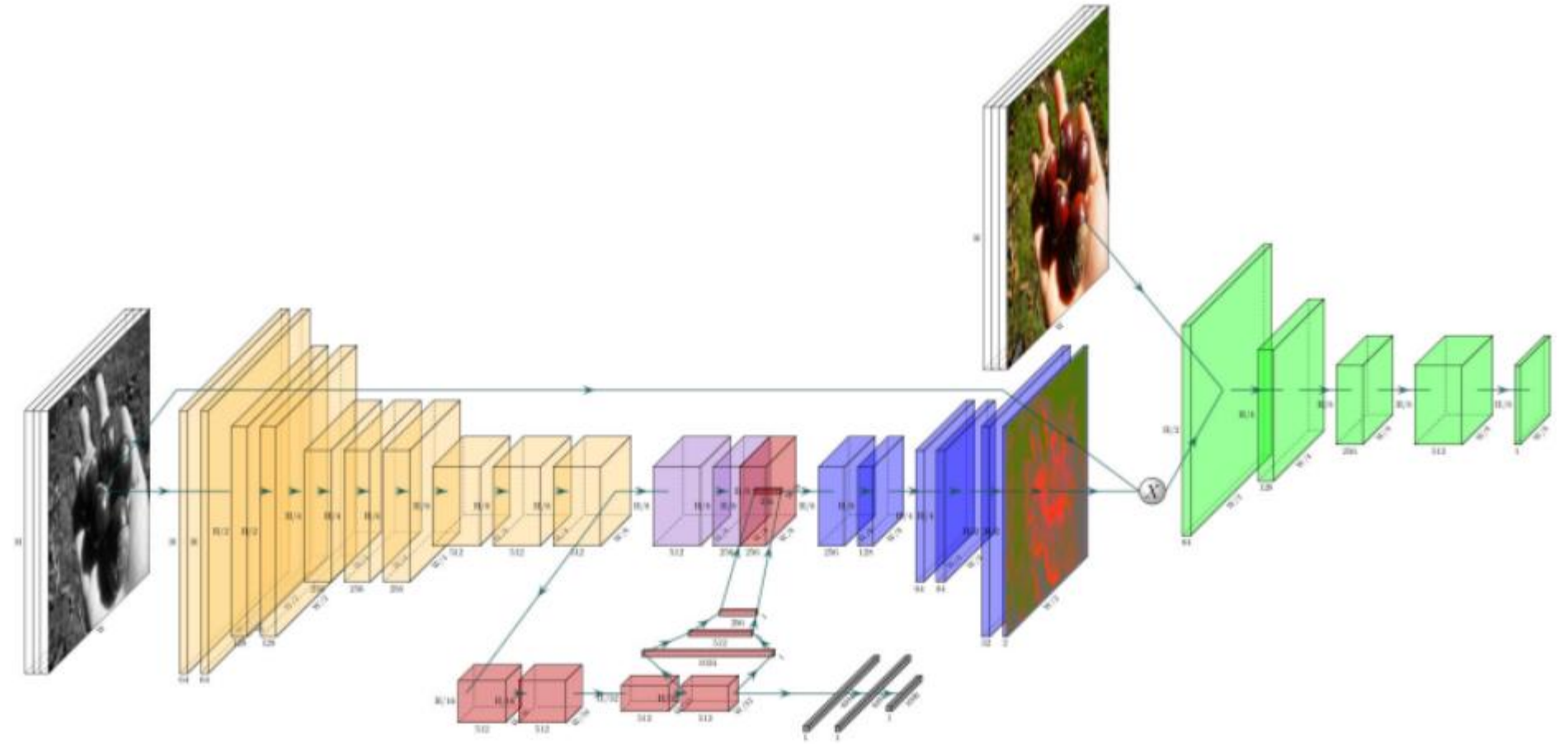
ChromaGAN의 특징으로는 같은 범주에 속하는 일부 물체를 다르게 색칠함으로써 가변성을 보여줌

SOTA(State-of-the-Art)

ChromaGAN _ 추가 설명

Discriminator network(Dn/초록색)

Generator network(G θ /노란색,보라색,빨간색,파란색)



Generator network는 G1 θ , G2 θ 인 두 개의 하위 네트워크로 구성

G1 θ (노란색,보라색,빨간색,파란색)는 색채정보 $(a, b) = G1\theta(L)$ 를 출력하고,

G2 θ (노란색, 빨간색, 회색)은 Class distribution vector, $y = G2\theta(L)$ 를 출력

SOTA(State-of-the-Art)

ChromaGAN _ 추가 설명

데이터셋

ImageNet에서 가져온 130만 이미지로 학습을 진행
흑백사진을 포함한 1000개의 다른 카테고리에서 다양한 색상을 학습
Fully connected layers를 사용하여서 Input size가 고정되어야 함
224 × 224 pixels를 VGG-16에서 학습

세부 모델

Epochs는 5를 사용하였고 Batch size는 10으로 설정

SOTA(State-of-the-Art)

DeOldify

Colorization은 영상물에 대한 semantic understanding을 기반으로 이뤄지는 task여서 일반적인 generative model만으로 좋은 성능을 기대하기 어려움

DeOldify는 self attention GAN에서의 self attention block을 추가하여 이미지 내 object들의 structural patterns을 학습하여 semantic understanding을 가능하게 해서 모델의 성능을 올림

SOTA(State-of-the-Art)

InstColor

Paper : Instance-aware Image Colorization, CVPR 2020

instance-aware colorization 방법 제안

: Object와 background를 명확히 나누어 colorization을 수행하는 방법

미리 학습된 object detector를 사용하여 object를 찾고 찾은 object에 대해 colorization 진행
-> object에 대해 colorization + 이미지 전체에 대한 colorization과 결합

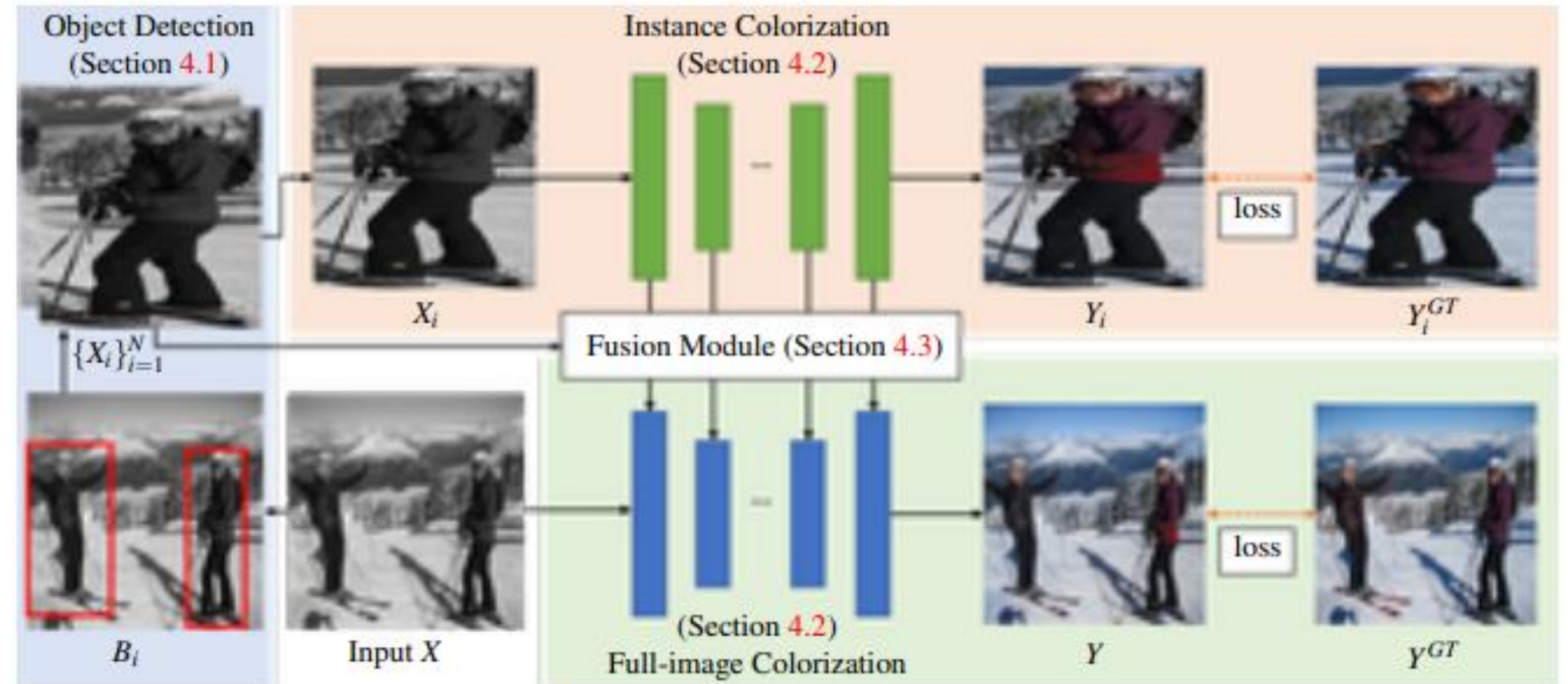
학습방법

- 1) 사전 학습된 detection모델로 object instance를 검출,
이를 통해 cropped된 object의 이미지를 만듦
- 2) 이미지 전체에 대한 네트워크와 object에 대한 네트워크 각각 학습
- 3) 2에서 학습한 두 네트워크에서 추출한 feature map을 Fusion

SOTA(State-of-the-Art)

InstColor _ 추가 설명

이미지 전체에 대해 Colorization을 하는 네트워크와 Object에 대해 Colorization을 하는 네트워크가 어떻게 동작 되는지에 대한 이미지



우선 입력으로 주어진 grayscale 이미지 X 가 주어졌을 때,
모델은 사전학습된 detection model을 이용해 X 에 대해서 bounding box(B_i)를 찾기 시작

그 다음 탐지된 모든 object(X_i)를 잘라내고 Instance network에서 X_i 를 colorization 하기 위해 사용

SOTA(State-of-the-Art)

InstColor _ 추가 설명

Fusion module을 사용하여
추출된 전체 이미지 feature map과 모든 layer에서의 Instance feature map을 fusion하는 것을 추천
(Instance의 색상이 예측된 배경의 색과 양립할 수 없을 가능성이 있기 때문에)

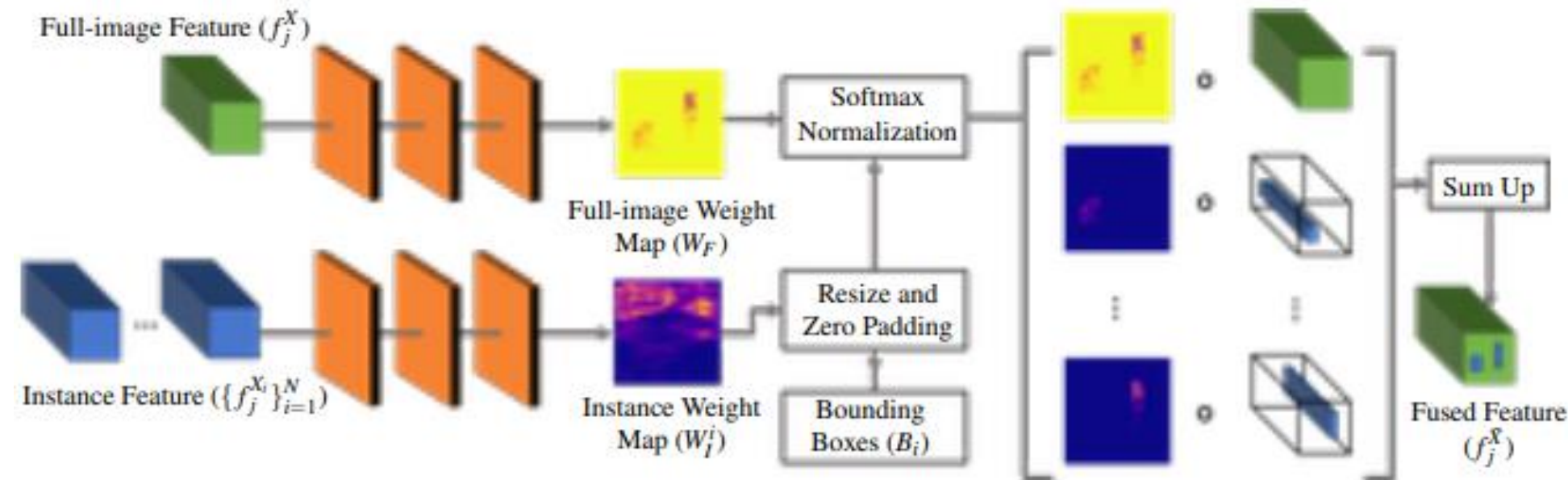
Fusion module

Fusion module의 input은 2가지

1. 전체 이미지에 대해 학습하는 네트워크로 부터의 feature(f_jX)
2. object 대해 학습하는 네트워크로부터 feature(f_jX_i)와 그에 대응하는 bounding boxes(B_i)

SOTA(State-of-the-Art)

InstColor _ 추가 설명



Fusion module에서 하는 일

- 3개의 convolutional layers로 구성된 작은 network를 통해 W_F 와 W_F^i 를 예측
- 그 다음 $f_j^{X_i}$ 와 $f_j^{X_i}$ 의 weight map(W_F^i)를 f_j^X 과 f_j^X 의 weight map(W_F)의 사이즈에 맞게 resize (이 때, padding은 0을 사용, resize된 $f_j^{X_i}$ 와 W_F^i 를 각각 $f_j^{X_i}$ W_F^i 라고 함.)
- 그 후 모든 weight map을 쌓고, 각각의 픽셀에 대해 softmax를 적용한 뒤, weighted sum을 하여 fused feature를 얻을 수 있음 (N : instance의 수)

실험 방식

1. 그냥 SOTA 모델에 한국 전통 이미지에 적용했을 때 잘 작동되는지 확인해보기
2. 데이터셋(한국 관련 이미지)으로 pre-train된 SOTA 모델들을 fine-tuning
3. 그 중 어느 모델이 더 나은지 실험, 비교해보기

Dataset

AI Hub

- 한국 이미지(음식)



구축 수량 : 150종 X 1,000장

AI Hub

- 한국형 사물 이미지

-> 이 중에서 궁궐, 가옥 등의 유적 건조물 이용

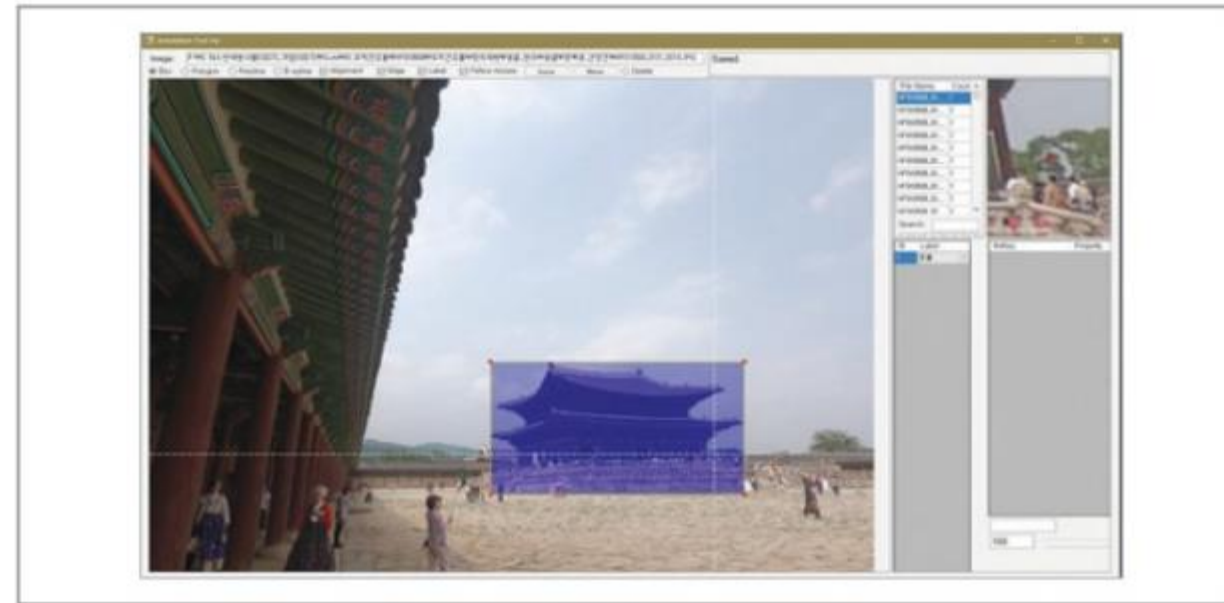


그림3 | 효율적인 데이터 구축을 위한 이미지 어노테이션 툴

유적건조물 250만장 (궁궐,가옥,탑,무덤 등)

삼품 80만장

랜드마크 20만장

Dataset

Dataset 구성하기

1. 해당 컬러 데이터 셋들을 흑백 이미지로 변환하여 한 쌍으로 구성되도록 하기
2. 각 데이터 크기를 256X256로 통일 시키기

평가 Metric

정량적 평가 (Quantitative Analysis)

: 숫자로 측정하고 표현하는 분석. 객관적으로 매겨질 수 있는 평가

정성적 평가 (Qualitative Analysis)

: 숫자가 아니라 말로 설명하는 분석. 인간의 주관적인 판단으로 매겨질 수 있는 평가

기본적으로 SOTA모델을 활용하기 때문에 정성적 평가로 주관적으로 나누기는 애매할 것 같아서 정량적 평가 사용

평가 Metric_정량적 평가

PSNR(Peak Signal-to-Noise Ratio)

: 실제 이미지와 생성된 이미지 간의 손실 정도를 판단하기 위한 평가 지표

FID(Frechet Inception Distance)

: 실제 이미지와 생성된 이미지 간의 특징 거리 측정에 가장 널리 사용되는 metric

SSIM(Structural Similarity Index Measure)

: 두 이미지 간의 유사성을 측정하는데 사용되는 metric

IQM(Image Quality Measurement)

: 생성된 컬러 이미지의 다채로움, 선명함, 대비를 평가하는 지표

출처

Colorization이란?

Towards Vivid and Diverse Image Colorization with Generative Color Prior 논문 사진 [https://arxiv.org/abs/2108.08826]

Colorization의 한계

이미지 복원 사진

[https://blog.naver.com/PostView.naver?blogId=baek2sm&logNo=222425129956&categoryNo=0&parentCategoryNo=0&viewDate=¤tPage=1&postListTopCurrentPage=1&from=postView]

SOTA(State-of-the-Art)

ChromaGAN - 논문 [https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=79]

InstColor - 논문 [https://arxiv.org/pdf/2005.10825v1.pdf]

Dataset

한국 이미지(음식) - AI Hub [https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=79]

한국형 사물 이미지 - AI Hub 데이터 설명 pdf [https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=144]

감사합니다!